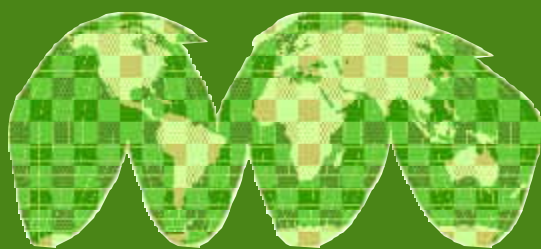


### **National Center for Education Statistics Disclaimer**

The information and opinions published here are the product of the International Indicators of Education Systems project's Network A and do not necessarily represent the policy or views of the Department of Education or the National Center for Education Statistics.



# Review of Assessment Activities

Issue 16

February/March 2004

## In this Issue

Hello to our friends and colleagues in the INES project! This first newsletter of 2004 presents information on how countries have been using data from PISA (the Programme for International Assessment) to analyze important policy issues and questions. The article provides information both on those results that have garnered the most media attention, as well as those that have spawned additional analysis, and it explores what impacts PISA may have had on policy and research in education in participating countries. It also provides some information on different national options in which countries are participating in PISA 2003, and their plans for disseminating results.

Also included in this issue is a country highlight focusing on assessment and testing in England. The article provides a wealth of information on the high-stakes examination programs in England, including the somewhat newer key stage exams; the use of school performance tables; and participation in international assessment activities. The article also discusses some of the main issues and debates that arise around measuring and documenting pupil performance in England. As usual, the newsletter also provides updates on Networks A, B, and C, and the PISA Board of Participating Countries, and a brief look at what is currently happening in national assessment programs in member countries.

We thank all those who contributed to the newsletter, including Jason Tarsh, from the Department for Education and Skills for contributing the article on pupil assessment and testing in England, and Jaap Scheerens and Maria Hendriks of the Netherlands for sharing information on Network C. We appreciate your efforts in keeping us informed of activities from around the INES Project. We hope you enjoy the latest newsletter!

## Learning from PISA

Last year, data collection for the second cycle of PISA (the Programme for Student Assessment), the OECD's study of the reading, mathematical and scientific literacy of 15 year-old students, was completed. Before results are released from PISA 2003 in December of this year, Network A thought it would be beneficial to examine the reception to and emerging impacts of the results from PISA 2000. So, we asked member countries about how PISA data are being used in their respective countries, what they are learning about their education systems, and what is happening in policy and practice as a result. We also asked them about their plans for dissemination of the 2003 results and if they had made any national enhancements to their participation in PISA on which they would be reporting. This article summarizes the responses from the 10 responding countries: Australia, Belgium-French community, Czech Republic, Denmark, England, Finland, the Netherlands, Slovak Republic, Sweden and the United States.

## Attention-getting results

One of the questions we posed to the newsletter countries was: which results have garnered the most attention? Because the results of analyses of PISA data differ across countries and across different measures, those results which get attention may, of course, also differ and may be related either to positive news or to areas which are in need of improvement, or still to the extent to which the results were expected or not. This was the case with the newsletter countries, although some trends were observed.

For example, one set of results that received much attention in multiple countries was the correlation between students' socio-economic background and their performance, with students with lower social background performing relatively less well than students whose families have more socio-economic resources. These findings have been discussed in **Australia, England, and Denmark**, where this relationship was found to be strong; and in the latter country, this finding has been a primary topic of conversation. In Denmark, officials were unpleasantly surprised by the results that showed that the social background of Danish students, to a higher extent, determines social outcomes than in other Nordic countries or in other top-performing countries like Canada, Japan and Korea. The respondent noted the im-

### Media Meter

The response from our English correspondent included this quick and interesting analysis, done by OECD, which shows the amount of press coverage across some of the PISA countries approximately one month after the release of the PISA 2000 data.

<u>Countries</u>	<u>No. of pages</u>
Germany	687
Switzerland	149
Canada	93
United Kingdom	88
Japan	84
Australia	54
Spain	53
United States	36
Belgium	32
New Zealand	25
Korea	21
Italy	16
Finland	8

plication of the finding: that Danish schools are less efficient in compensating negative social heritage than we desire keeping in mind that 'equity' and 'fair chances for all' are core social values in Denmark.

Other results that received attention in multiple countries also related to differences between different groups of students. For **Australia** and **Belgium (French community)**, wide variation between the best and weakest students have raised policy concerns. Similarly, the low variation PISA showed among Finnish students has drawn positive attention there. In **Finland**, however, the primary concern is the wide differences favoring girls in reading literacy. Despite their overall strong

performance and low variation, Finland has one of the largest gender gaps of all OECD countries, and policy makers are highly concerned about stimulating the interest, engagement and self-confidence of boys in reading, which PISA also showed were correlated with performance. The underachievement of boys is of particular concern, not only in Finland but in **Australia**, as well. Other student differences countries are watching with interest and concern are between students of differing racial or ethnic backgrounds—for example, Indigenous students in Australia and immigrant students and those with a foreign background in **Sweden**.

In the **Czech Republic**, the differential performance of students in reading literacy versus the other domains (with the former being below the international average) has been a key story. **Sweden** also interestingly noted that some of the findings that generated the most interest there originated outside of Sweden. For instance, it was a Norwegian report that included measures of discipline in schools in Nordic countries that sparked numerous news stories and articles on the conditions in Sweden and action from concerned parents, who called and emailed the National Agency for Education in record numbers.

### Preliminary impacts

Although the full effects of PISA on policy, practice, or research will not be seen for some time, some of the countries responding to our newsletter did offer some preliminary reports on the impacts PISA 2000 is having on these areas. **Australia** was very positive about the importance of PISA, which was described as being “invaluable in improving understanding of middle school achievement” and helpful in corroborating minimum benchmark standards in literacy and numeracy in Year 7. In fact, the federal and state/territory ministers have agreed that PISA will be a key source for informing progress against the *National Goals for Schooling for the 21<sup>st</sup> Century*.

Both Belgium (French community) and Denmark described how the results of PISA 2000 helped to reinforce the previously emerging trends to focus more in their education systems on the quality of outcomes. In **Belgium (French community)**, shortly after the release of PISA 2000, an important law was passed with respect to educational monitoring. Although, as our respondent pointed out, the law was prepared prior to PISA, the comparative information on French Belgian students helped strengthen the argument for its necessity and turned around some reluctant stake-holders. **Denmark**, similarly, noted that PISA has contributed significantly to the focus on the quality of outcomes. Here, PISA was seen as fitting in to a broad shift, also evidenced by recent reforms to the curriculum frameworks for primary and lower secondary education, which introduced more national involvement in this previously highly decentralized system. In both these cases, concern with results—either comparatively lower average performance or relatively sizeable variation between students—may have helped support the institutionalization of changes already underway.

In a contrasting, though not totally contradictory, situation, **the Netherlands** noted the relatively lesser activity following the release of their national report of the PISA 2000 results compared with the 1991 IEA Reading Literacy Study, in which a surprisingly low performance

### Edu-Tourism?

One unique contribution of PISA is that, in Finland at least, it has led to a new type of tourism, reports Pirjo Linnakylä of the Institute for Educational Research at the University of Jyväskylä. “During the last two years Finnish schools, the Ministry of Education, research institutes, and teacher training colleges have received hundreds of visitors from different countries—particularly Germany, Japan, Korea, Sweden, Denmark, Austria, Switzerland, and England. It has been very interesting to listen to their stories, how they see our school system and instructional practice and how they interpret the reasons behind Finland’s success in PISA. Some look for explanation in the tradition of oral storytelling and the national epic, *Kalevala*; others from the small class sizes or teachers’ autonomy and competence handling heterogeneous groups in the comprehensive school; and still others from the cold and dark winter nights, forgetting that summer nights are light and warm.”

spurred a host of policy actions such as curricular reform and teacher training.

For at least two countries (and clearly more, based on the information presented in the next section), PISA has led to new areas of study. The **French community of Belgium** is studying the relationship of primary school and libraries, hypothesizing that increased library resources could help reinforce early competencies in reading. In **Sweden**, PISA results prompted the minister to call for an in-depth analysis of the performance and outcomes for students with a foreign background. **Finland** noted that PISA's impact—outside the confirmation of the good (e.g., high performance) and more concerning news (e.g., the gender gap) and the perhaps less obvious impacts (see Box 2)—has been in raising awareness and questions around research and methodological innovations. For one, there is an interest in understanding the link between international and national assessments. Also, there is a growing interest in looking for socio-cultural explanations for results and employing sophisticated multi-level modeling techniques in analyzing student outcomes.

Like Finland, **Australia** also described an interest in the linkage of PISA and (in this case) state-level assessment. In the State of Western Australia, some of the publicly released PISA items have been pilot tested with locally developed items. It is hoped that these items will scale with local items and give Western Australia the capacity to link its historical reading scale to the PISA scale. This effort, it is further hoped, would allow PISA to penetrate closer to classroom practice and, if successful, would be followed by similar efforts with the mathematics and science items in the future.

### Publications from PISA 2000

Many of the countries that participated in PISA 2000 prepared their own reports providing overviews of the results from the national perspective. These countries included Australia, Austria, the Flemish and French communities of Bel-

gium, Brazil, Canada, Czech Republic, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, New Zealand, Norway, Portugal, Sweden, Switzerland, the United Kingdom, and the United States. A comprehensive list of national reports, along with links to the electronic form of the reports or ordering information, is available from the PISA website. (The address for this site and the sites of the other reports mentioned in this section are summarized in table 1.)

Some countries also prepared follow-up reports, analyzing particular issues of interest (such as previously described) in more depth. Here, we limit ourselves to reports from newsletter respondent countries. For example, **Australia** reported that several thematic reports are currently in preparation, including reports on the performance of Indigenous students, the impact of geographic location on student performance, immigrant status and student performance, and language background and student performance, as well as a report that compares the results of PISA and TIMSS. These reports are expected to be available from the national center for PISA at ACER. Australia also noted how PISA had been reported in at least two other government reports, including one on teaching and teacher education and one on mapping Australia's science and innovation system.

**Sweden** recently completed an in-depth analysis of the results for students with a foreign background and published a report on the topic, with an executive summary in English forthcoming this year. Discussions also have begun about additional secondary analyses, with possible issues to explore being connections between computer use and interest and reading literacy, classroom factors and reading literacy, and learning strategies and reading literacy from a Swedish perspective. Sweden is also making use, as other countries do, of the existing international publications—translating and adapting reports or report chapters of interest. For example, officials are translating the chapter “Improving both quality and equity: Insights from PISA 2000” from OECD's *Education Policy*



**Table 1: Resources**

All countries	<ul style="list-style-type: none"> <li>• <a href="http://www.pisa.oecd.org/NatReports/cntry.htm">http://www.pisa.oecd.org/NatReports/cntry.htm</a></li> </ul>
Regional reports	<ul style="list-style-type: none"> <li>• Northern Lights on PISA (<a href="http://www.pisa.oecd.org/NatReports/PISA2000/NorwayNorthernLights.pdf">http://www.pisa.oecd.org/NatReports/PISA2000/NorwayNorthernLights.pdf</a>)</li> </ul>
Australia	<ul style="list-style-type: none"> <li>• National report and forthcoming thematic reports and TIMSS-PISA comparisons (<a href="http://www.acer.edu.au">http://www.acer.edu.au</a>)</li> <li>• Review of Teaching and Teacher Education (<a href="http://www.dest.gov.au/schools/teachingreview/documents/Main_Report.pdf">http://www.dest.gov.au/schools/teachingreview/documents/Main_Report.pdf</a>)</li> <li>• Mapping Australia's Science and Innovation System (<a href="http://www.dest.gov/mapping/main_report.htm">http://www.dest.gov/mapping/main_report.htm</a>)</li> </ul>
Czech Republic	<ul style="list-style-type: none"> <li>• Tasks for Measuring Reading, Mathematical and Science Literacy</li> <li>• Knowledge and Skills for Life: Reading, Mathematical, and Science Literacy of 15 year-old students in OECD countries</li> <li>• Results of Czech Students in International Surveys</li> <li>• Non-Traditional Items or We Read with Understanding</li> <li>• <a href="http://www.uiv.cz">http://www.uiv.cz</a></li> </ul>
Belgium (French)	<ul style="list-style-type: none"> <li>• Most recent report is No. 13 and 14 of <i>Les Cahiers du SPE</i> (<a href="http://www.ulg.ac.be/pedaexpe/cahiers.html">http://www.ulg.ac.be/pedaexpe/cahiers.html</a>)</li> </ul>
Denmark	<ul style="list-style-type: none"> <li>• Education Peer Effects: Quantile Regression Evidence from Denmark with PISA 2000 Data (<a href="http://www.pisa.oecd.org/NatReports/PISA2000/DenmarkEducationPeerEffects.pdf">http://www.pisa.oecd.org/NatReports/PISA2000/DenmarkEducationPeerEffects.pdf</a>)</li> </ul>
Finland	<ul style="list-style-type: none"> <li>• Well prepared for the future: PISA 2000 in Finland (<a href="http://www.jyu.fi/ktl/pisa/PISA-RAPORTTIscreen.pdf">http://www.jyu.fi/ktl/pisa/PISA-RAPORTTIscreen.pdf</a>)</li> <li>• The Finnish success in PISA – and some reasons behind it (<a href="http://www.jyu.fi/ktl/publication1.pdf">http://www.jyu.fi/ktl/publication1.pdf</a>)</li> </ul>
Sweden	<ul style="list-style-type: none"> <li>• A report on the situation of students with a foreign background (<a href="http://www.skolverket.se">http://www.skolverket.se</a>)</li> </ul>
United States	<ul style="list-style-type: none"> <li>• Outcomes of Learning: Results from the 2000 Program for International Assessment of 15 Year-Olds in Reading, Mathematics and Science Literacy</li> <li>• Highlights from the 2000 Program for International Student Assessment</li> <li>• A Comparison of the National Assessment of Educational Progress, the Third International Mathematics and Science Study Repeat and the Program for International Student Assessment</li> <li>• Comparative Indicators of Education in the United States and Other G-8 Countries: 2002</li> <li>• <a href="http://www.nces.ed.gov/surveys/pisa/Publications.asp">http://www.nces.ed.gov/surveys/pisa/Publications.asp</a></li> </ul>

*Analysis* 2002 and working on Swedish-focused reports based on the PISA international thematic reports on self-regulated learning and on engagement in school.

In the **Czech Republic**, at least two of the follow-up reports have focused on the assessment tasks, providing examples of items, examples of Czech students' responses, and information for teachers in working with these type of items and with literacy assessments. In **Denmark**, researchers used sophisticated techniques to explore the issue of peer effects in education. Again, information on these reports and those from other

newsletter countries are noted in the resources table at the end of this article.

Also, in countries with strong regional divisions, reports focusing on the individual regions were prepared. For example, in **Belgium**, researchers in both the Flemish and French communities worked with PISA data (which, internationally, was reported for Belgium as a whole) to produce reports tailored specifically for their respective communities. Similarly, in the **United Kingdom**, there are separate reports for England, Northern Ireland, and Scotland.

Finally, our newsletter respondents noted some interesting regional reports, including two reports examining results for the Nordic countries: *Northern Lights on PISA*, published in 2002, and Youngsters' reader profiles (*Nuorten lukijaprofileja*), which is in press. Seven education systems also are participating in a further comparative study being led by the German Institute for International Educational Research.

## Plans for 2003

### National options

We also asked our respondents if their countries were participating in any national options in PISA 2003 in order to be able to provide additional information on topics, populations, or questions of interest. Similar to PISA 2000, most countries do make some enhancements at the national level and these enhancements generally relate to oversampling, sampling additional populations, or supplementing the questionnaires.

Australia, Slovak Republic and Sweden each noted that they would be oversampling the population in order to allow for reporting results disaggregated for regions or groups of interest. For example, **Australia** oversamples students in the smaller states and territories, as well as Indigenous students, who comprise only 3 to 4 percent of the population but about whom information on outcomes is needed since improving the education of these students is a high national education priority. Similarly, **Sweden** oversamples schools with a high proportion of students with a foreign background in order to allow for analyses that may inform educational improvements for this growing group of students. **Slovak Republic** oversamples students in order to be able to report results not only for the country's 8 geographic regions but also for the 7 school tracks, or education programs.

Two countries also sample additional populations to receive the PISA assessments. The **Czech Republic** samples students who are in the 9<sup>th</sup> grade

but who are not 15 to give a more comprehensive picture of performance in the country and its regions at this level of schooling. **Denmark** adds a sample of 16 year-olds. Because Danish children begin schooling comparatively late, policy makers are interested in the comparative outcomes of the one-year older students, as well.

Three countries also reported adding components (or taking up international optional components) to student questionnaires. Both the **Czech** and **Slovak Republics** asked students questions about their educational careers, and both the Czech Republic and **Finland** added items on ICT usage or literacy. Finland also asked about students' engagement and diversity in reading, which was powerful predictor of achievement, and the Czech Republic added items about family background.

Finally, in PISA 2003, **Australia** also sampled 12,000 students to form the 2003 cohort of the on-going Longitudinal Surveys of Australian Youth (LSAY). LSAY is a longitudinal study of the progress of young people between school, post-secondary education and work. The use of PISA to establish cohorts under LSAY allows for more comprehensive information on student achievement to be collected and linked with students' progress beyond schooling into further education and employment. Consideration also is being given to establishing further LSAY cohorts in future PISA cycles.

### Reporting and dissemination

Looking ahead to the release of PISA 2003 data and results, most of the respondents indicated that their countries would be releasing a national report or summary on the same day of the international release. This includes Australia, Belgium (French community), England, Finland, Slovak Republic, Sweden, and the United States. At least one country, **Czech Republic**, will participate in a national press conference and comment on the international release date but will delay release of a full report until later a later date. **Finland**, which will release a short report of about

100 pages on the international release date, plans to finalize a more lengthy report for release at a national conference in the second quarter of 2004.

While most countries are still in the process of making specific or detailed plans for dissemination and secondary analyses for PISA 2003, some countries did share their preliminary plans. For example, the **Slovak Republic**—which is going through this for the first time, not having participated in PISA 2000—plans to follow up the national report, which will focus on introducing the PISA literacy concept and the Slovak educational system outcomes from the international perspective, with at least two additional publications. The follow-up reports, expected in 2005 and 2006, will address (1) results by region and school track and (2) the literacy domains and task/item format and type—the latter for informing innovation in teacher preparation and for didactic applications, as well as possible curricular reform aimed at strengthening general education in each type of school at the ISCED 3 level. A report for Czech

Republic, Hungary, Poland and the Slovak Republic also is under consideration. Another example of a country with advanced plans is **Sweden**, which has a variety of ideas about possible products from PISA 2003. These include a report comparing PISA, TIMSS and national tests; a handbook for reading and understanding the national report with sections targeted to different stakeholders (e.g., teachers, students, school heads); a second Northern Lights study; a report on the items (including p-values, framework classification, etc.); a trend report for reading and mathematical literacy; and articles in Swedish professional journals.

In **Denmark**, although exact plans are uncertain, our newsletter respondent noted that the release of the PISA 2003 results will be closely coordinated with those of the Pilot Review to Examine Quality and Equity in School Outcomes, a project being conducted in conjunction with the OECD. A strategy to plan for dissemination around both these activities is being prepared this summer.

## Network Updates

### Network A

Network A last met on October 21-22 in Lisbon, Portugal. At the meeting, members discussed Network A's long-term data strategy; reviewed the indicators proposed for *Education at a Glance (EAG) 2004*; discussed the work of the Network A/C Task Force on Teaching and Learning; and received a report on the outcomes of the ICT literacy assessment feasibility study.

With regard to the long-term data strategy, the discussion focused both on issues that relate broadly to a data strategy for INES and those that relate more specifically to the future of PISA. Members discussed such questions as: what data do we need to be collecting in the future, what skills will students need, and what target populations are of greatest interest? The Network A Secretariat is redrafting a formal

strategy for the Network, in accordance with the conversations during the plenary meeting, which emphasized an outcome-driven nature, explaining how different activities fit together in a bigger picture, and not neglecting components such as communication and reporting, and in accordance with developments in the BPC's Strategic Development Group.

Members also discussed the draft text and figures for the Network's chapter on learner outcomes in *EAG 2004*, which included three new indicators (i.e., trends in reading literacy in 4<sup>th</sup> grade, 15 year-old students' engagement in school, and engagement and academic performance) and three repeat indicators (i.e., reading literacy in 4<sup>th</sup> grade in 2001, and reading literacy and mathematical and scientific literacy of 15 year-olds in 2000). This combination of new and



repeat indicators was both to be more responsive to the desire at OECD to move toward providing a more comprehensive picture of competencies in each edition of *EAG* and to cover the different work areas in the INES conceptual framework related to Network A.

At this meeting, members also welcomed Irwin Kirsch from the Educational Testing Service (ETS), the head of the ICT expert group, who gave a presentation of the outcomes of the study to assess the feasibility of administering an assessment of ICT literacy in PISA using the newly developed framework and pilot platform and instruments. He described the goals of the study, the materials developed for use in the study, how the feasibility study was conducted, and finally what the results and lessons learned were. Although the Network was very complimentary of the quality of the work, especially under the tight timeline ETS and its partners were working under, it was later determined by the BPC that the cost of full implementation of the assessment in 2006 was prohibitively high.

Finally, the Network reviewed the draft strategy paper of the Task Force on Teaching and Learning, which presented a three-pronged strategy for collecting relevant information for INES: (1) an expansion of Network C's existing indicators; (2) a teacher survey across ISCED levels on teachers' attitudes, characteristics and perceptions; and (3) additional development work to elaborate a more long-term plan for collecting information on the impacts of teaching.

The next meeting of Network A will be in Lucerne, Switzerland on March 18-19, preceded by a meeting of the BPC on March 15-17 and tentatively to be followed by a meeting of the

Task Force on Teaching and Learning on March 22-23.

## Network B

Network B last met on February 3-5, 2003 in Madrid, Spain. Eighteen countries were represented,

as well as the OECD Secretariat and delegates from Eurostat. As we reported in the last newsletter, at this meeting, members discussed issues related both to *EAG* indicators and to development work ongoing in five key areas: educational attainment; continuing education and training (CET); equity, including social outcomes; transition from education

to work; and rates of return to education.

Of note, members reviewed a report on the development of a CET module, which is to serve as a set of guidelines for the development of internationally comparable indicators on CET based on data collected via household surveys. Countries were invited to submit comments on the open issues in the report. Network B also discussed the development work on Young Adults with Low Levels of Education (YALLE), a data collection effort coordinated by the Swedish Secretariat that surveyed 20 to 24 year-olds who have not attained ISCED level 3 and who are enrolled neither in an education nor in a work-study program. Based on results from the pilot study in 8 countries, members had decided at their previous meeting to continue with YALLE and recommended that it should be incorporated in regular OECD data collections. At this meeting, it was decided that Sweden would conduct a new data collection using the pilot data collection as a template. Finally, stemming from a presentation on PISA-L, members

### Did you know... ?

You can find information about Network A's membership and activities, as well as meeting records dating back to 1994 and all previous issues of the Network A newsletter on the NCES website at: <http://www.nces.ed.gov/surveys/international/INES/>. Check it out! And, if you have any comments or suggestions, please don't hesitate to let us know ([mstephens@air.org](mailto:mstephens@air.org)).

decided that a small task group would work on suggestions for harmonizing school leaver surveys and longitudinal surveys.

Network B will next meet in Canberra, Australia on February 9-11, 2004. At this meeting, members will review progress in the different development areas, as well as review an overall strategy paper for Network B and specific workplan for 2004. Specifically with regard to CET, members will be updated on countries' comments to the report on the CET module and will review the revised report and recommendations for data collection and possible inclusion in *EAG 2005*. Members also will discuss how to measure and present information related to the equity dimension of labor market and social outcomes; potential future indicators from the YALLE study and on social outcomes of education, as well as proposed indicators for *EAG 2004*; and updates on work related to transitions from school to work.

### Network C

In the period July to December 2003, Network C's main activities have focused on finalizing the report from the International Survey of Upper Secondary Schools (ISUSS), planning and data collection for indicators for *EAG 2004*, and reviewing the work related to the Network A/C Task Force on Teaching and Learning. Network C last met in Korea on November 18-20, 2003.

At this meeting, Network C members discussed the outcomes of the INES priority rating exercise for the activities within the remit of the Network, as well as the implications for its work, and it was agreed that the Network should prepare a formal program of work for the next 3 to 4 years. This program of work will be structured according substance (e.g., indicators on the system, school, and classroom teacher level, with reference to malleable and antecedent conditions) and main data sources (e.g., national statistics/panels, surveys, secondary analysis of international data sets, and—in the long term—perhaps observation studies). The Network C

Secretariat will produce a first draft of the plan early in 2004 and send it to members for comment.

The Network also decided to postpone the international launch of the ISUSS report, initially planned for July 2003, until January 2004. This decision was taken in order to expand the sections on policy conclusions and to arrange an effective launch and publicity. In the meantime, countries which had participated in ISUSS began working on individual national reports, and Denmark, Norway, Finland and Sweden collectively prepared a Nordic report.

Also at the meeting, members discussed the content of Chapter D of *EAG 2004*. Rather than a succession of indicators, members thought that the Network C chapter should attempt to paint more of a joint picture of its theme. The Network endorsed the proposals for the indicators on decision-making and the three core Network indicators (teachers' working and teaching time, instructional time, and salaries). Additionally, the Network proposed to prepare an indicator based on data from ISUSS. New indicators for 2004 will relate to admission/grouping/placement policies and feedback from stakeholders. With regard to indicators on decision-making, the Network decided to focus on changes in the profile of decision-making (level and mode) between 1998 and 2003, and to first analyze those aspects that experienced changes. With regard to the core Network C indicators, the Network was asked to reconsider their definitions in the light of the definitions proposed in the OECD Handbook for Internationally Comparative Statistics of Education, and the Eurydice data and definitions on teacher working conditions and pay.

Regarding the work of the Task Force, Network C members provided a general endorsement of the overall strategy paper and endorsed further preparatory work in both the areas of attitudes and expectations and teaching effectiveness. The Network had a lengthy discussion on the proposal for a teacher survey, facilitated through

organization into working groups on the topic. One of the main concerns among Network C members was that, while a cross-ISCED survey was highly desirable, a survey at ISCED level 2 was thought to be more feasible. A second main concern was related to the scope of the survey and whether or not it should be limited to perceptions and attitudes. Members of Network C supported a broader survey including information on teaching practices and strategies and on teachers' background characteristics, and offered specific suggestions in this vein. To proceed, the Network C Secretariat will improve the substance of the proposal for the survey on teacher attitudes and expectations, starting with the policy areas the activity is intended to inform and presenting discussion on pros and cons of different options relating to the ISCED level and links to PISA and the strategy more broadly. In doing this, the Network C Secretariat also will ask advice of the people working at the CERI project *Attracting, Developing and Retaining Effective Teachers*.

The next meeting of the INES Network C will be in Greece (dates to be confirmed).

### Board of Participating Countries

The PISA Board of Participating Countries (BPC) last met in Lisbon, Portugal in October 2003, in conjunction with the Network A meeting. The meeting agenda was quite full, with items related to PISA 2000, 2003, and 2006. Related to PISA 2000, members reviewed the recently published thematic reports and those in preparation, established plans for further analytic work, and discussed how to facilitate multi-lateral cooperation in the analysis of PISA results. With regard to PISA 2003, the current cycle, the BPC discussed establishing a strategy for the development of thematic reports,

reviewed the implementation of the assessment and compliance with technical standards, and established a policy for the public release of PISA 2003 assessment material. Related to PISA 2006 and beyond, members reviewed progress with the development of the PISA 2006 science literacy framework, finalized the contractual arrangements for PISA 2006, and decided on future perspectives for the integration of ICT components into PISA, as well as on the applications of non-OECD countries to participate in PISA 2006—15 of which were accepted at the time. With regard specifically to the ICT work, members decided that the costs for implementation of the proposal would be prohibitively high and instead requested a new call for tender for an ICT component, which would have a reduced scope of work that would focus on adding value to the assessment of science knowledge and skills.

The BPC also elected several new officers for the chair and vice-chair positions and created the position of Special Advisor, individuals who also will collectively serve as the PISA Executive Group. Ryo Watanabe from Japan was elected as the new chair of the BPC, and Dianne Pennock from Canada and Anita Wester from Sweden were confirmed as vice-chairs. Gerard Bonnet from France also was nominated for the next vice-chair position available. And finally, in recognition of his contributions and leadership, the outgoing chair, Eugene Owen, was invited to serve as the Special Advisor to the BPC, a role in which he will provide counsel on long-term strategic development and new work areas, as well as guidance on maintaining the scientific integrity, operational feasibility, and sustainability of PISA over time.

The next BPC meeting will be held in Lucerne, Switzerland on March 15-17, 2004.

## Country Highlight: Pupil Assessment in England

Prepared by Jason Tarsh

*This article is, I hope, an informative, broad-brush account of the situation regarding assessment and the curriculum in England, which also seeks to capture the flavour of some of the related debates and research. It is not offered as an official or definitive description, and any opinions that are expressed are my own views and should not be taken as necessarily representing those of the Department for Education and Skills (DfES).*

In education, the UK has never been a unitary state, and especially not at the school level. Scotland, joining in 1707 with the Act of Union, has always had its own system and Northern Ireland, created in 1921 following independence for Southern Ireland, still retains fully selective secondary schools. So, this article will focus on England although, with recent increased devolution of government—including power over education—to Wales, Scotland and Northern Ireland, the UK now shows some striking internal contrasts in education policy, including in pupil assessment, and the final section describes these briefly.

### The English system in outline

England has a population of around 49 million and a school population of around 8 million, with the average size of each age year around 600,000 pupils. (And it is pupils we would never say students, which we reserve for post-16s). There are around 24,000 schools in England. In the state (*maintained*) sector there are roughly

18,000 primary and 3,500 secondary schools (the 21,500 total including some 400 middle schools), plus about 2,500 *independent* (private but, confusingly also known as *public*) schools that can be primary, secondary or all-through. The independent sector is small: currently about 8 percent of pupils aged 11-15 are in these schools (though they account for a rather higher proportion of 16-18 year olds, “sixth-formers”)—a proportion that has remained pretty much stable over the 1990s. These schools receive no government subsidy<sup>1</sup> and necessarily cater to well-to-do families. They also are, at secondary level, a high-attaining sector in terms of pupil attainment, as our own exams and PISA showed.

Primary education starts at age 5<sup>2</sup> and carries on to age 11 when all pupils then transfer to secondary schools which can be either 11-18 or 11-16. We have just a few middle schools. In the latter case, the pupils who “stay on”—16 being the end of compulsory schooling—transfer to the sixth form of an 11-18 school, a Sixth Form College or a Further Education College.<sup>3</sup>

Almost all secondary schools in England are comprehensive, having changed, starting in the mid-1960s under the then Labour government, from a previously tripartite system<sup>4</sup> based on selection tests at age 11 (the *11-plus*).<sup>5</sup> There are, additionally, some 164 *grammar schools*, which select pupils at age 11, still using the *11-plus*, and which now account for about 4 per-

<sup>1</sup> This is in contrast to, for example, France where fees at private *lycées* are heavily government subsidised. Independent schools in England, however, do receive certain tax benefits as charitable institutions.

<sup>2</sup> The English terminology refers to *Years* rather than grades so age 5/6 is Year 1, age 6/7 Year 2 through to age 10/11, which is Year 6, the end of primary. The first year of secondary school is then Year 7 and ends at age 15/16, which is Year 11. Thus, the TIMSS 14 year-olds were in grade 8 but English Year 9, and the PISA sample covered Years 10 and 11.

<sup>3</sup> Sixth-form and further education colleges are administered and funded as a separate sector from the schools and consist of some 397 further education colleges and 45 sixth-form colleges.

<sup>4</sup> The three parts were: grammar schools, secondary modern schools and technical colleges. The technical colleges were always a small, residual sector although the intention had been that they would be a significant third, vocational sector.



cent of secondary pupils. It is local education authorities (some 150 in England) and part of elected local government bodies (councils) that determine whether their schools are comprehensive or selective. Ten LEAs have chosen to remain fully selective,<sup>6</sup> with a fair number of other LEAs having one or a few grammar schools in an otherwise comprehensive system.

In the 1960s, 70s and much of the 80s, Education Department officials would speak of the schools as a “national system, locally delivered” with the centre having little power and the schools being run largely by LEAs. Assessment was similarly decentralised. There were two high-stakes national pupil assessments: GCE (General Certificate of Education) Ordinary level (*O-levels*) taken at age 16, supplemented in the mid-1960s by CSEs,<sup>7</sup> and GCE Advanced level (*A-levels*) taken at age 18/19. For both, the syllabi and grading were determined by private *examination boards*. The boards were originally created by consortia of universities and had a strong regional focus (e.g., one covered the north and midlands, another covered the south) and tended to be the main examining body for the schools within their respective regions, although schools could use any board’s exams if they wished.

The Conservative (Margaret Thatcher) governments of the 1980s introduced radical and wholesale change, most notably with major legislation in 1988, which meant central government taking much more power over the schools and LEAs while requiring LEAs to devolve much of their previous power and control over resources to individual schools. The most no-

table of the reforms were the introduction of a statutory national curriculum from ages 5 to 16 with four *Key Stages* (5-7, 7-11, 11-14 and 14-16) and the introduction of *Key Stage Tests* at ages 7, 11 and 14, in addition to the already re-organised GCSE (General Certificate of Secondary Education, which had replaced the previously separate O-levels and CSEs) at age 16.

## Pupil Examinations

### *O-levels/GCSEs and A-levels*

Both these high-stakes exams have existed since the late 1940s, albeit with O-Levels being transformed into GCSEs, as described above. The exams are taken in individual subjects and a pupil gets a public grade in each (in contrast to the French and German grouped exams with pass or fail and unpublished marks). The GCSEs run on a scale from A\*, A, and B down to G, and A-levels range from A to E.<sup>8</sup> Pupils typically take around seven GCSEs and those who go on to A-level would typically take three subjects, but some take just two and others four or more.

Pupils have some choice after age 14 in which individual subjects they study, with the National Curriculum subjects of English, maths, science and ICT being compulsory. However, they have complete freedom in which individual subjects and which combinations of subjects they take GCSE exams, although almost all pupils would take maths and English GCSE. At A-level the system has long ceased to be specialised, and there really is very little limit to the diversity of subjects with, for example, French, biology and geography entirely feasible. In practice, since A-levels are the main academic qualification for

<sup>5</sup> The 11-plus tests consisted essentially of tests of IQ and basic skills, with the content decided locally.

<sup>6</sup> As with the German *Länder*, there are marked variations in the proportions of pupils in the grammar schools in the fully selective areas, ranging from 26 to 46 percent.

<sup>7</sup> O-levels were seen as catering for the grammar school stream, the top 30 percent or so of the ability range. CSEs were created in the mid-1960s to meet a demand from the increasing

numbers of secondary modern pupils wanting to acquire qualifications.

<sup>8</sup> In certain subjects, such as maths, there are three “tiers” for curriculum, exams and grades and, for example, the lowest tier in maths would not cover algebra. This approach is intended to cater to differences in ability and to give less able pupils an incentive to tackle a difficult subject. The highest grade available on the middle tier is B and on the lower, D, so they do overlap. Pupil grades would not show which tier they had taken.

entry to a degree course (although, unlike for example France and Germany, they do not confer a right of entry and equally a university can waive any such entry requirement), pupils planning to go on certain degree subjects will take certain subjects (maths, physics and chemistry being a long-established choice). But there are many degree courses, especially in the “new” social sciences (psychology, business studies, accountancy), as well as in established subjects like law, where there are typically no subject requirements.

While GCSEs and A-levels are national exams, they are, as noted above, designed and set by private, examination boards. These boards are, however, overseen by a DfES “Non-Departmental Public Body” agency, the Qualifications and Curriculum Authority (QCA), as part of the 1988 reforms. Starting in the mid-1980s, the boards started to compete nationally for schools to take up their exams and, since they are, to a limited degree, commercial organisations, they have an incentive to expand their market share. This might be one cause of the great variety of exam subjects and subject syllabi that there now seems to be.

Based on take-up, the most widespread government-led recent innovation at 14-18 has been the introduction of General National Vocational Qualifications (GNVQs) at three levels, the upper two essentially matching GCSE and A-levels. These are in vocational subjects such as business studies, tourism, etc. Much more recently, and as part of a continuing concern over providing high status vocational qualifications, the government has introduced vocational GCSEs.

<sup>9</sup> In 1998 the Department also introduced statutory baseline assessment for all new 4 or 5 year- old pupils. This is for individual diagnostic purposes and there are no national tests but all of the 90 or so different tests currently in use must be centrally approved.

<sup>10</sup> The official tables publish the results of the schools in alphabetical order and so are not ranked.

## Key Stage Exams

Briefly, the current system has six main levels of attainment, numbered simply 1 to 6. Each year all 7, 11 and 14 year-olds are tested in maths, English and science (although there are no formal science tests at age 7) with additional, optional subjects at age 14. In addition to the written tests, which are externally marked, each child also is assessed by their teacher in the three subjects and so has two grades, which need not be the same.

DfES has set specific levels which it expects pupils to reach: Level 2 and above at KS 1 (age 7), Level 4 and above at KS 2 (age 11, the end of primary school) and Level 5/6 at KS 3 (age 14). When the Levels were first set, they were explicitly aspirational and, for example, fewer than half of 11 year-olds reached the expected Level 4 when the first KS 2 exams took place in 1995. Table 2 shows overall trends in key stage exams over time.

As well as the statutory Key Stage tests, there are also additional, optional tests for use by schools to assess pupil progress within Key Stages and which most pupils take. More recently, and as part of its *Key Stage 3 Strategy* (ages 11-14), DfES has introduced further such optional tests including a *Year 7 Progress Test* in English and maths, which is aimed at pupils who did not meet the expected standard at the end of KS 2.<sup>9</sup>

## School Performance Tables

Closely entwined with all our national tests are what are officially *school performance tables* but which are usually referred to in the media as the “school league tables.”<sup>10</sup> Launched in 1994, and then just for secondary schools, these now show, for each named primary and secondary school, the proportions of their pupils achieving the various Key Stage Levels at KS 2 and 3 (but not 1) and, at secondary level, the proportions achieving various combinations of numbers and grades at GCSE and A-level.

**Table 1: Percentage of primary pupils achieving the expected levels in Key Stage 1 (age 7) and Key Stage 2 (age 11), 1995-2003**

	1995	1996	1997	1998	1999	2000	2001	2002	2003
<b>KS 1</b>									
Reading	78	78	80	80	82	83	84	84	84
Maths	78	82	84	84	87	90	91	90	90
<b>KS 2</b>									
English	49	57	63	65	71	75	75	75	75
Maths	45	54	62	59*	69	72	71	73	73
Science	70	62	69	69	78	85	87	86	87

\* A new, mental arithmetic test was introduced in this year.

Note: KS 1 tests were introduced in 1991, KS 2 in 1995 and KS3 (not shown) in 1993. The new (Labour) government in 1997 set KS 2 targets of 80 percent achieving at or above Level 4 in English and 75 percent in maths by 2002. These targets were not met. There are now revised targets of 85 percent at or above Level 4 and 35 percent at Level 5 by 2004 for both English and maths.

There have been periodic calls for the tables to move to a value-added basis, that is, to allow for pupils' prior attainment, initially by using a proxy measure for social disadvantage such as Free School Meals eligibility since true longitudinal prior attainment measures have only recently become available. Following extensive development work, DfES published the first value-added tables for secondary schools in 2002 and for primary schools in 2003. A second round of KS 2 to KS 3 and KS 3 to GCSE tables have been published even more recently.

### International assessments

England, whether as England or the UK, has taken part in all the recent international pupil assessments, with the exception of the 1991 IEA Reading Literacy Study.<sup>11</sup> This apparent desire in England for international comparative information seems to go back a long time, indeed back to Victorian times and now as then, probably partly stems from long-running concerns about the country's international economic competitiveness.

In both TIMSS 1995 and 1999, England performed well in science but was only average in maths, and thus was one of a sizeable group of

countries with contrasting performance in the two subjects. This maths-science divide also held for primary maths, which would normally be taught by the same science teacher, leading some to suggest that curricular factors might have played a part in the results. However, it is also possible that primary teachers are more at ease with teaching science than maths. In the two more recent literacy studies, PISA and PIRLS, England did pleasingly well although there are no precedents to judge how far we might have progressed nor indeed to suggest why we did so well in reading.

The various international assessment results certainly had major press coverage on the day of their release and, in PISA and PIRLS, this was very much celebratory rather than looking for bad news. However, at the same time, there also was little attempt to suggest or analyse why we had performed as we did. In TIMSS 1995, the focus was very much on the only average performance in maths and the particular weakness in arithmetic, with the science ranking probably under-played. As so often, the press coverage was probably closely influenced by what was immediately accessible to the general reader, and shortcomings in arithmetic are naturally newsworthy. Academic interest in the studies has been disappointingly limited and has not prompted much in the way of debate and further analysis. The

<sup>11</sup> See the Times Education Supplement, 16 October 1992, which discusses possible and contrastig reasons why England did not take part (available from author, J. Tarsh).

UK academic community does, though, embrace two leading critics of such studies!

## The public and policy debate over assessment

Assessment and curriculum issues have proved to be frequent sources of public debate and controversy in England. That is probably an intrinsic feature of these topics, and it is difficult to say whether the debates are more intense here than elsewhere. What does seem clear is that, in a centralised and quite centrally directed system, national government and its agencies are always directly implicated and the debate is necessarily a national one. Furthermore, there is a lot of documentary and statistical information to debate about! There is, some might say, also still a continuing echo in the current debates of the controversies sparked off by the educational radicalism of the 1960s.

Briefly, the long-standing and recurring issues have tended to be over:

- “*Standards.*” One persistently debated question is how far the trends in pupil attainment recorded by statistics are genuine and how far they instead reflect the impact of, for example, changes in curriculum content, changes in methods of assessment (e.g., coursework, accounting for 20 percent of the marks in GCSEs; modular assessment; and the use of contextualised questions proving contentious), and teaching to the test.<sup>12</sup> The debate here is fairly continuous but also tends to follow a cycle with a peak just after the annual publication of the GCSE

<sup>12</sup> There have been various studies of “standards” over time but none that has been definitive enough to end the debate. See, for example, *The Guardian*, 18 December 2003, “Doubt cast on primary pupils’ progress,” for a press account of a very recent official research report on primary standards. Another, closely related QCA study included a very interesting account of public, academic and teacher perceptions of whether “standards” had changed over time. See supplemental table for details.

and A-level results each August. Table 3 charts the long term progress of GCSE attainment in England.

- *Subject breadth, particularly at ages 16-18.* Some in the education community are looking to make post-16 study of some subjects, such as maths, compulsory or to move to an *English baccalaureate* where pupils have to take a wider range of subjects than now, albeit with some as minors.<sup>13</sup>
- *Specific subjects in the curriculum.* Other major issues are, for example, whether pupils should have to study foreign languages or science post-16, what the role of studying Shakespeare in the English GCSE should be, and recently, whether pupils’ choices in history A-level were overly dominated by study of World War II.
- *The burden of testing, on schools, teachers and on pupils.* It has recently been argued, without real documentary evidence, that English

**Table 2: Proportion of school-leavers/15-year-olds getting five or more GCSE A\*-Cs or equivalent, selected years, 1963 to 2003**

Year	Percent	Year	Percent	Year	Percent
1963	16	1985	27	2001	50
1965	20	1988	30	2002	52
1970	23	1990	35	2003	53
1975	23	1995	44		
1980	24	2000	49		

NOTES: 1998 was the first year of the GCSE examinations. Beginning in 1988/89, the statistics switched from a basis of school-leavers (of any age) to 15 year-olds.

SOURCE: DfES.

The relatively flat trend between 1965 and 1987 has been attributed by some commentators to a policy of tacit norm-referencing by the examination boards although others have denied that this took place. Interestingly, A-level grades and the proportions of higher degree classes also took off in the 1980s although perhaps not in the expected order, with degree class rising first, followed by A-level, and then GCSE!

<sup>13</sup> On this issue, some recently published research sponsored by QCA, which sought to compare the demands of A-levels and the International Baccalaureate, concluded that, insofar as they could be compared, one was not obviously more demanding than the other. See, for example, *The Guardian*, 29 December 2003, “A-level tests knowledge better than bac.” There are also AS levels, deemed to be of A-level standard but involving a shorter period of study. These were introduced as a means to promote greater subject breadth, but they have had only limited take-up by pupils.



pupils are “the most tested in the world” and that the KS 1 tests are unacceptably stressful for 7 year-olds. The government has reaffirmed its view that “testing is here to stay” and that tests are a means to securing higher pupil attainment and to addressing the needs of the socially disadvantaged.

- *School performance tables.* These tables have been a continual source of dispute for, let it simply be said, the usual reasons that seem to occur in any country where this is a live issue. The recent publication of the value-added tables seems not to have ended the controversy despite many within the schools having previously called for just this method.

Three other issues deserve mention although they are not so high profile.

The widespread use of national examinations in employee selection is a testament to their continued economic significance but, at the same time, there has also been a boom in the use of other forms of employee selection by psychometric tests and the like. Some recent research (e.g., Wolf and Jenkins, see the supplemental table) found that employers were giving increasing weight to psychometric tests and this reflected a perceived greater need to make the right choice of recruit but also some loss of faith in academic qualifications.<sup>14</sup>

In the past, commentators have noted what seemed to be inappropriate use of exams in employee selection and some researchers have concluded that what employers (and university selectors) are looking for from testing is ever finer rankings within the conventional elite groups

rather than seeing exams and grades as ways of identifying talent across a wider range of ability. This has been reflected in the assessment system with, in the early 90s, the splitting of the GCSE “A” grade into A\* and A. There also have been more recent calls from some universities for the A-level A grade to be split yet again. It is relevant here that the standard system-level performance indicator used in many official publications (and echoed in most other commentary on school performance) is the proportion of pupils gaining “five or more GCSE A\* to Cs.”<sup>15</sup>

Finally, England has recently introduced a fully longitudinal system of tracing pupils through the school system using a unique pupil identifier. The resulting *Pupil Level Annual Schools Census* (PLASC) will allow much better analysis than previously of pupil progression and hence the impact of attainment at one age on attainment later. One limitation though is the lack of information on pupils’ personal and family circumstances with the main measure being eligibility for Free School Meals, a rough proxy for low income. We are, however, hoping as an ad hoc exercise, to draw on the much more extensive SES information on the pupil sample in PISA 2000 by linking this to the pupils’ Key Stage 3 results.

## Comparisons within the UK

It might be thought that with four constituent territories, there would be scope for useful comparative education work within the UK. In practice this has not been the case, probably partly reflecting England being by far the largest of the four so that it might not seem to be useful

<sup>14</sup> There is often ambiguity in employers’ views on the role of qualifications and, by implication, the schooling they are certifying. Thus a very recent statement by the CBI, our main employers’ organisation, commenting on proposed reforms to 16-19 assessment says: “Employers recruit an individual person, not primarily a qualification. School leavers’ qualifications may not necessarily reflect the competencies required for work. Many of the qualities employers look for when appointing a young person are difficult to measure in terms of qualifications – for example, personality, attitude and

flexibility. Factors such as interest, enthusiasm, willingness to learn and ability to do the job take precedence over qualifications in employers’ recruitment criteria.” See: <http://www.cbi.org.uk/ndbs/press.nsf/0363c1f07c6ca12a8025671c00381cc7/1181c5a7e963413280256e270054e996?OpenDocument>

<sup>15</sup> It is debatable whether GCSE grades below C, and certainly below D, have any real currency either in employment or in access to further education.

comparator after all. Furthermore, there have not been marked differences in attainment within the UK as far as can be judged.<sup>16</sup> Perhaps illustrating all this and also the strength of feeling that assessment issues can arouse, following the recently introduced government devolution, the new Wales and Northern Ireland assemblies have decided to end the publication of school league tables and also to end testing of primary pupils.<sup>17</sup> The Scottish Executive is currently considering taking the same route. So we now have different parts of the same country reaching sharply contrasting policy views on these matters—a new light, this, on learning from others!

## Stop Press

Two very recent major DfES commissioned reports by external bodies on curriculum and assessment have just been published and are useful guides to many of the current debates and possible future directions of policy. The interim report of the Working Group on 14-19 Reform (the “Tomlinson Review”) and the Smith report on school maths. The Smith report includes detailed discussion of GCSE and A-levels and employers’ views of qualifications. See the Supplemental Table for URLs for the downloadable reports.

## Current Assessment Activities

Among the countries that responded to our request for information, several countries described activities for national assessments related to development, data collection, and reporting that occurred between June and December 2003 and currently are ongoing.

For example, assessment-related development and test construction activities include the following examples.

- In **Australia**, work is progressing on national sample assessments in primary science, and in information and communication technology skills and civics and citizenship at both the primary and secondary levels. The first national sample survey of primary school science was conducted in the second half of 2003 and will be conducted every three years. Results of this assessment are expected to be available in the first half of 2004. The first national sample civics and citizenship assess-

ment will take place in late 2004 and ICT in late 2005. These will complement assessments at the state/territory level, which are described in the next section.

- In the **Slovak Republic**, test construction is occurring for two important measures of learning outcomes: (1) the national assessment of the knowledge and skills of students in the 9<sup>th</sup> grade of compulsory education in Slovak language and mathematics, which will occur this month; and (2) the *Maturita* examinations, which are the leaving exams for the end of secondary education.
- In **Sweden**, test construction is an ongoing activity, undertaken by university institutions, for the national testing program. This program is based on the national curriculum and syllabi and is intended to support teachers’ evaluations of students’ strength and weaknesses and to evaluate the extent to

<sup>16</sup> Scotland and Northern Ireland have traditionally outperformed England at GCSE although, for Scotland, this depends on an assumed equivalence between Scottish Standard grades (their GCSE-counterpart) and the English GCSEs. In the international assessments, England performed better than Scotland in PIRLS 2001 and better in TIMSS

1995 at age 14 in science (but not in maths), although there was virtually no difference in scores in PISA 2000.

<sup>17</sup> A new interim report for the Welsh Assembly (the Daugherty report) is now proposing also ending external tests at KS 3. See [http://www.learning.wales.gov.uk/scripts/fe/news\\_details.asp?NewsID=1123](http://www.learning.wales.gov.uk/scripts/fe/news_details.asp?NewsID=1123).

which students have reached national goals and hence ensure fair and equal grading throughout the country. Test construction activities include item construction and piloting, data collection and analysis, development of scoring instructions, and standard setting. The timing of the tests are in the spring for 11 and 15 year-olds and both autumn and spring for student in primary school.

- Also in **Sweden**, the National Agency for Education recently released an item bank of materials to support teachers' assessment of students. The bank includes materials for subjects in the national tests, including Swedish, English, and mathematics, as well as other subjects such as French, German, Spanish, physics, biology, and vocational subjects. The items in the bank can be used for diagnostic as well as summative purposes, and complements the national tests in the overall process of assessing learning.

Data collection and reporting activities included the following.

- In **Australia**, each state and territory conducts assessments of student performance in literacy and numeracy at years 3, 5, and 7 on an annual basis. This testing, which usually occurs sometime in the second half of the year, occurred in August in 2003. The results of these assessments are "equated" so that nationally comparable results can be reported. This yearly *National Report on Schooling in Australia* can be accessed at <http://www.curriculum.edu.au/anr/index.html>. However, it should be noted that the latest version available is the report covering the year 2000.
- In 2003, **Belgium (French community)** tested 7<sup>th</sup>-grade (secondary) students on their competencies in sciences. The paper-and-pencil test consisted of both multiple-choice and open-ended items. This effort is part of the French community of Belgium's voluntary national testing program, which is organized for one grade level and one subject each year. The purpose of the tests, which most schools participate in, is to provide teachers with formative information—the tests are given at the beginning of the school year—about the level of achievement and knowledge of their students. Between June and the present, related activities included final preparation of the tests, administration of the tests in schools, marking of a representative sample of tests by researchers, report preparation, and the development of didactical suggestions for teachers (e.g., in the form of brochures, booklets, and other documents).
- **Denmark** administered its annual national examinations (which include oral and written components) in 9<sup>th</sup> and 10<sup>th</sup> form in June 2003. The results of these examinations are reported to the Ministry of Education, including information on the average grades for each school. The latter statistic is owing to a government initiative to provide public information on "raw" scores—that is, without any modifications relating, for example, to the socio-economic intake of the school—a policy which has been subject to debate. Denmark also plans to participate in OECD's Pilot Review to Examine Quality and Equity in School Outcomes, which will involve in-depth analysis and review team visits around the topic of outcomes and related activities.
- In **Finland**, the National Board of Education carried out assessment of students' achievement in mother tongue and literature (both in Finnish and Swedish) in the 6<sup>th</sup> grade of basic comprehensive school. This sample-based national assessment found that, among other results, there was variation between girls in boys in reading, linguistic skills and writing – though overall average performance was good. Variation in teachers' grading practices also was revealed.
- In the **Netherlands**, students in the last year of primary school are preparing for examinations, which are occurring this month and

which will determine to which level of secondary school they will go. Secondary school students will take centralized examinations later in the year in May and June.

- The **Slovak Republic** is currently analyzing the results from its 2003 assessment of 9<sup>th</sup> grade students' knowledge and skills in the Slovak language and in mathematics.
- Data collection for several activities was underway in the **United States** roughly during this period. The National Assessment of Educational Progress' (NAEP) Technology Rich Environments (TRE) study, which ended in May 2003, examined 8<sup>th</sup>-grade students' abilities to explore and synthesize scientific information online. This was a complement to two previous components,

including writing online and mathematics online. Beginning in October 2003 and continuing through May 2004, NAEP will be conducting its long-term trend assessments in reading and mathematics for 9, 13, and 17 year-olds. Later this year, NAEP also will pilot a foreign language assessment in Spanish in approximately 200 high schools throughout the country.

- The **United States** also released several reports during the period, including the results of the NAEP 2003 reading and mathematics assessments for the 4<sup>th</sup> and 8<sup>th</sup> grades, as well as the results of a the Trial Urban District Assessments in reading and mathematics, which was conducted in 9 large urban districts throughout the country.

#### Supplemental Table on Resources

Entity	Web address and notes
Department for Education and Skills (DfES)	<ul style="list-style-type: none"> <li>• <a href="http://dfes.gov.uk/index.htm">http://dfes.gov.uk/index.htm</a> (home page)</li> <li>• <a href="http://www.standards.dfes.gov.uk/performance">http://www.standards.dfes.gov.uk/performance</a> (how schools can use assessment for school improvement)</li> <li>• <a href="http://www.standards.dfes.gov.uk">http://www.standards.dfes.gov.uk</a> (information for parents)</li> <li>• <a href="http://www.dfes.gov.uk/performance/tables">http://www.dfes.gov.uk/performance/tables</a> (data)</li> <li>• <a href="http://www.dfes.gov.uk/primarydocument">http://www.dfes.gov.uk/primarydocument</a> ("Excellence and Enjoyment," publication from July 2003 that sets out the Department's primary strategy and launches a new, three-level target setting regime for primary schools)</li> <li>• <a href="http://www.14-19reform.gov.uk">http://www.14-19reform.gov.uk</a> (Tomlinson review)</li> <li>• <a href="http://www.mathsinquiry.org.uk/index.html">http://www.mathsinquiry.org.uk/index.html</a> (Smith report)</li> </ul>
Qualifications and Curriculum Authority	<ul style="list-style-type: none"> <li>• <a href="http://www.qca.org.uk">http://www.qca.org.uk</a> (home page of agency responsible for overseeing curriculum and assessment; note: few research publications currently available on-line)</li> <li>• <a href="http://www.qca.org.uk/products/95_6300.html">http://www.qca.org.uk/products/95_6300.html</a> (URL for the "Massey" study of standards over time)</li> </ul>
Assessment Reform Group (ARG) and related	<ul style="list-style-type: none"> <li>• <a href="http://www.assessment-reform-group.org.uk">http://www.assessment-reform-group.org.uk</a> (home page of group of English academics who promote the use of formative, versus summative, testing)</li> <li>• <a href="http://www.edweek.org/ew/ewstory.cfm?slug=27blackbox.h21&amp;keywords=Olson">http://www.edweek.org/ew/ewstory.cfm?slug=27blackbox.h21&amp;keywords=Olson</a> (article on ARG)</li> <li>• <a href="http://eppi.ioe.ac.uk/EPPiWeb/home.aspx?page=reel/reviews.htm#DfES">http://eppi.ioe.ac.uk/EPPiWeb/home.aspx?page=reel/reviews.htm#DfES</a> (recent "systematic review" of the impact of summative assessment and tests on students' motivation for learning)</li> </ul>
Critics	<ul style="list-style-type: none"> <li>• <a href="http://www.ioe.ac.uk/hgpersonal/">http://www.ioe.ac.uk/hgpersonal/</a> (web-site for Harvey Goldstein, well-known education academic and persistent critic of school league tables and international assessments)</li> </ul>
Centre for the Economics of Education	<ul style="list-style-type: none"> <li>• <a href="http://cee.lse.ac.uk/index.html">http://cee.lse.ac.uk/index.html</a> (home page)</li> <li>• <a href="http://cee.lse.ac.uk/publications.htm">http://cee.lse.ac.uk/publications.htm</a> (articles by Wolf and Jenkins on employer use of selection tests and implications for academic qualifications: DP12, DP27 and DP 29; also by Wolf: Growth Stocks and Lemons: Diplomas in the English Market Place, <i>Assessment in Education</i>, 4(1) 1997)</li> </ul>
Other	Collection of recent press articles and related published material on some of the issues discussed in the article available from the author ( <a href="mailto:jason.tarsh@dfes.gsi.gov.uk">jason.tarsh@dfes.gsi.gov.uk</a> ).

This newsletter is published under the auspices of Network A. Network A, which is primarily concerned with indicators of learner outcomes, is one of several working groups that are part of OECD's international Indicators of Education Systems (INES) Project. The newsletter is prepared by Eugene Owen (Network A Chair) and Maria Stephens and Bela Shah of the American Institutes for Research with contributions from Network A members. The newsletter was designed by Charmaine Llagas.



## Network A Representatives

Mr. Eugene Owen (*Chair*)  
International Studies Program  
National Center for Education  
Statistics  
UNITED STATES

Ms. Wendy Whitham  
Department of Education, Science  
and Training  
AUSTRALIA

Mr. Helmut Bachmann  
Federal Ministry of Education,  
Science and Culture  
AUSTRIA

Mr. Jürgen Horschinegg  
Federal Ministry of Education,  
Science and Culture  
AUSTRIA

Ms. Christiane Blondin  
Service de Pédagogie expérimentale  
de l'Université de Liège  
BELGIUM

Mr. Luc Van de Poele  
Department of Education  
Universiteit Gent  
BELGIUM

Mr. Michael O'Gorman,  
Curriculum & Assessment Policy  
Branch  
Ministry of Education Ontario  
CANADA

Mr. Jerry Mussio  
Centre for Education Statistics  
Statistics Canada  
CANADA

Ms. Pavla Zieleniecova  
Institute for Information on Education  
CZECH REPUBLIC

Mr. Jørn Skovsgaard  
Ministry of Education  
DENMARK

Ms. Pirjo Linnakylä  
Institute for Educational Research  
University of Jyväskylä  
FINLAND

Mr. Thierry Rocher  
Ministère de l'éducation nationale  
FRANCE

Mr. Jochen Schweitzer  
Der Senator für Bildung und  
Wissenschaft  
Der Freien Hansestadt Bremen  
GERMANY

Ms. Evangelia Varnava-Skoura  
Ministry of Education  
GREECE

Ms. Zsu Zsa Hamori-Vaczy  
Ministry of Education  
HUNGARY

Mr. Gerry Shiel  
Educational Research Centre  
St. Patrick's College  
IRELAND

Mr. Ryo Watanabe  
Department of International  
Education  
National Institute for Educational  
Research  
JAPAN

Mr. Myungjoon Lee  
International Research and  
Cooperation Korea Institute of  
Curriculum and Evaluation  
KOREA

Ms. Iris Blanke  
Ministère de l'éducation nationale  
LUXEMBOURG

Mr. Fernando Córdova  
Secretaria de Educacion Publica  
MÉXICO

Mr. Paul van Oijen  
Ministry of Education Science and  
Research  
NETHERLANDS

Ms. Lynne Whitney  
Ministry of Education  
NEW ZEALAND

Ms. Anne-Berit Kavli  
Norwegian Board of Education  
NORWAY

Ms. Glória Ramalho  
Gabinete de Avaliação Educacional  
(GAVE)Ministério da Educação  
PORTUGAL

Mr. Vladislav Rosa  
Director of Central School Inspection  
SLOVAK REPUBLIC

Mr. Guillermo Gil  
Instituto Nacional de Calidad y  
Evaluación  
Ministerio de Educación y Ciencia  
SPAIN

Ms. Anna Barklund  
Swedish National Agency for  
Education  
SWEDEN

Ms. Anita Wester  
National Agency for Education  
SWEDEN

Mr. Erich Ramseier  
Amt fuer Bildungsforschung  
SWITZERLAND

Mr. Sevki Karaca  
Directorate for Research and  
Development  
TURKEY

Mr. Jason Tarsh  
Department for Education and Skills  
UNITED KINGDOM

Ms. Elois Scott  
International Studies Program  
National Center for Education  
Statistics  
UNITED STATES